

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
22 August 2002 (22.08.2002)

PCT

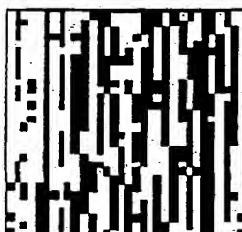
(10) International Publication Number
WO 02/065782 A1

- (51) International Patent Classification⁷: H04N 7/26, A., C., M.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). BAGGEN, Constant, P., M., J.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). OOSTVEEN, Job, C.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
H04H 9/00
- (21) International Application Number: PCT/IB02/00379
- (22) International Filing Date: 7 February 2002 (07.02.2002) (74) Agent: SCHMITZ, Herman, J., R.; Internationaal Octrooibureau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (25) Filing Language: English
- (26) Publication Language: English (81) Designated States (*national*): CN, JP, KR.
- (30) Priority Data: (84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).
01200505.4 12 February 2001 (12.02.2001) EP
01202720.7 17 July 2001 (17.07.2001) EP
- (71) Applicant: KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). Published:
— with international search report
- (72) Inventors: HAITSMA, Jaap, A.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). KALKER, Antonius, For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: GENERATING AND MATCHING HASHES OF MULTIMEDIA CONTENT



51



52



53

(57) Abstract: Hashes are short summaries or signatures of data files which can be used to identify the file. Hashing multimedia content (audio, video, images) is difficult because the hash of original content and processed (e.g. compressed) content may differ significantly. The disclosed method generates robust hashes for multimedia content, for example, audio clips. The audio clip is divided (12) into successive (preferably overlapping) frames. For each frame, the frequency spectrum is divided (15) into bands. A robust property of each band (e.g. energy) is computed (16) and represented (17) by a respective hash bit. An audio clip is thus represented by a concatenation of binary hash words, one for each frame. To identify a possibly compressed audio signal, a block of hash words derived therefrom is matched by a computer (20) with a large database (21). Such matching strategies are also disclosed. In an advantageous embodiment, the extraction process also provides information (19) as to which of the hash bits are the least reliable. Flipping these bits considerably improves the speed and performance of the matching process.